Course guide
200612 - ADL - Longitudinal Data Analysis

Last modified: 12/06/2023

Unit in charge: School of Mathematics and Statistics
Teaching unit: 715 - EIO - Department of Statistics and Operations Research.
749 - MAT - Department of Mathematics.

Degree: MASTER'S DEGREE IN STATISTICS AND OPERATIONS RESEARCH (Syllabus 2013). (Optional subject).

Academic year: 2023 **ECTS Credits:** 5.0 **Languages:** English

LECTURER

Coordinating lecturer: CARLES SERRAT PIE

Others: Segon quadrimestre:
BERNAT PLANDOLIT LÓPEZ - A
CARLES SERRAT PIE - A

PRIOR SKILLS

The prior skills that are desirable are the ones from basic courses in mathematical statistics and probability in the degree courses. Two referencies that can help to prepare in this preliminary phase are:

Gómez, G. (2002) Estadística Matemàtica 1 (Teoria). Apunt de la FME. Universitat Politècnica de Catalunya.
Gómez, G, Nonell, R and Delicado, P. (2002) Estadística matemàtica 1. (Problemes). Apunts de la FME. Universitat Politècnica de Catalunya

It is supposed that the student knows the linear model and the generalized linear model. This knowledge can be previously obtained and consolidated in the subject on Linear Models that it is taught during the first semester.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
3. CE-1. Ability to design and manage the collection of information and coding, handling, storing and processing it.
4. CE-3. Ability to formulate, analyze and validate models applicable to practical problems. Ability to select the method and / or statistical or operations research technique more appropriate to apply this model to the situation or problem.
5. CE-4. Ability to use different inference procedures to answer questions, identifying the properties of different estimation methods and their advantages and disadvantages, tailored to a specific situation and a specific context.
6. CE-6. Ability to use appropriate software to perform the necessary calculations in solving a problem.
7. CE-7. Ability to understand statistical and operations research papers of an advanced level. Know the research procedures for both the production of new knowledge and its transmission.
8. CE-8. Ability to discuss the validity, scope and relevance of these solutions and be able to present and defend their conclusions.
9. CE-9. Ability to implement statistical and operations research algorithms.

Transversal:
1. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

2. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

TEACHING METHODOLOGY

The content, the learning method and the evaluation of this subject have been designed taking into account criteria of sustainability (in particular the environmental ones), social commitment (care to third parties, security or well-being) and gender perspective (treatment of cases and examples).

The course is practical and PBL oriented (Project / Problems Based Learning).

Specifically:

a) Outline the methodological needs from real data analysis,

b) Develop the theoretical model (interest will be focused on the modeling and interpretation of results and, secondarily, in demonstrating the theoretical results).

c) Return to the data to perform the analysis and interpretation of results.

Labs sessions will be in R.

LEARNING OBJECTIVES OF THE SUBJECT

Longitudinal data combine information from the variability between individuals and the evolution and variation within individuals. For this reason, they represent, by their frequency and relevance, a challenge not only for the professional statistician but also for the theoretical development.

The course objective is, first, to develop the theoretical framework and, second, to implement the knowledge gained by using the statistical software R.

STUDY LOAD

Type	Hours	Percentage
Hours small group	15,0	12.00
Hours large group	30,0	24.00
Self study	80,0	64.00

Total learning time: 125 h

CONTENTS

Linear Mixed Model (LMM).

Description:
Linear Mixed Model (LMM).

Full-or-part-time: 36h
Theory classes: 6h
Practical classes: 6h
Self study : 24h

Generalized Estimating Equations (GEE).

Description:
Generalized Estimating Equations (GEE).

Full-or-part-time: 25h
Theory classes: 4h 30m
Practical classes: 4h 30m
Self study : 16h

Generalized Linear Mixed Model (GLMM).

Description:
Generalized Linear Mixed Model (GLMM).

Full-or-part-time: 16h 40m
Theory classes: 3h
Practical classes: 3h
Self study : 10h 40m

Introduction to Missing Data Analysis.

Description:
Introduction to Missing Data Analysis.

Full-or-part-time: 33h 20m
Theory classes: 6h
Practical classes: 6h
Self study : 21h 20m

Extensions: Longitudinal Data Analysis with multivariate response and Joint Modeling.

Description:
Longitudinal Data Analysis with multivariate response and Joint Modeling.

Full-or-part-time: 14h
Theory classes: 3h
Practical classes: 3h
Self study : 8h

GRADING SYSTEM

- 20%: Homework to be done during the semester (report, presentation and defense). Task in group of 2-3 students.
- 10%: Report on a paper. Individual task delivered to the professor.
- 10%: Quiz in the Campus Digital (Atenea). Single answer multiple choice test and with penalization.
- 60%: Final exam (Theory -development questions: 30%, Laboratory -data analysis: 30%)

EXAMINATION RULES.

a) In the assessment of the Homework a 10% of self-assessment and peer assessment of the various groups will be taken into account.
b) Language for the Homework and the Report on a paper is English.
c) Final exam:
c1) In this first part of the exam (theory and modeling questions) the student can NOT have the course material, but only writing instruments and calculator.
c2) In the laboratory part the student may have all the course material (in paper and/or digital).

BIBLIOGRAPHY

Basic:
- McCulloch, C.E.; Searle, S.R. Generalized, linear and mixed models. New York: John Wiley & Sons, 2001. ISBN 047119364X.
- Molenberghs, G.; Verbeke, G. Models for discrete longitudinal data [on line]. Springer, 2005 [Consultation: 28/06/2023]. Available on: https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/0-387-28980-1. ISBN 0387251448.
- Verbeke, G.; Molenberghs, G. Linear mixed models for longitudinal data [on line]. Springer-Verlag, 2000 [Consultation: 28/06/2023]. Available on: https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/b98969. ISBN 0387950273.
- Little, Roderick J.A.; Rubin, D.B. Statistical analysis with missing data [on line]. 3rd ed. John Wiley & Sons, 2019 [Consultation: 28/06/2023]. Available on: https://onlinelibrary-wiley-com.recursos.biblioteca.upc.edu/doi/book/10.1002/9781119013563. ISBN 0471183865.

Complementary:
- Verbeke, Geert; Fieuws, Steffen; Molenberghs, Geert; Davidian, Marie. "The analysis of multivariate longitudinal data: A review". National Institute of Health-Public Access [on line]. [Consultation: 28/06/2023]. Available on: https://www.researchgate.net/publication/224811683_The_analysis_of_multivariate_longitudinal_data_A_review.- Faraway, Julian James. Extending the linear model with R : generalized linear, mixed effects and nonparametric regression models. Boca Raton (Mass.): Chapman & Hall/CRC, 2006. ISBN 9781584884248.
- McCullagh, P.; Nelder, J.A. Generalized linear models. 2nd ed. Chapman & Hall, 1989. ISBN 0412317605.
- Crowder, M.J.; Hand, D.J. Analysis of repeated measures. Chapman and Hall, 1990. ISBN 041231830X.
- Pinheiro, J.C.; Bates, D.M. Mixed effects models in S and S-Plus [on line]. Springer-Verlag, 2000 [Consultation: 28/06/2023]. Available on: https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/b98882. ISBN 0387989579.
- Schafer, J. Analysis of incomplete multivariate data. Chapman & Hall, 1997. ISBN 0412040611.
- Verbeke, G.; Molenberghs, G. Linear mixed models in practice a SAS-oriented approach. Springer-Verlag, 1997. ISBN 0387982221.
- Diggle, P.; Liang, K-Y.; Zeger, S.L. Analysis of longitudinal data. 2nd ed. Oxford University Press, 2002. ISBN 0198524846.
- Lindsey, James K. Models for repeated measurements. 2nd ed. Clarendon Press, 1999. ISBN 0198505590.
- Galecki, Andrzej; Burzykowski, Tomasz. Linear mixed-effects models using R : a step-by-step approach [on line]. New York: Springer, 2013 [Consultation: 28/06/2023]. Available on: https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/978-1-4614-3900-4. ISBN 9781461438991.
- Rizopoulos, Dimitris. Joint models for longitudinal and time-to-event data : with applications in R [on line]. Boca Raton, FL [etc.]: Chapman and Hall/CRC, cop. 2012 [Consultation: 28/06/2023]. Available on: https://www-taylorfrancis-com.recursos.biblioteca.upc.edu/books/mono/10.1201/b12208/joint-models-longitudinal-time-event-data-dimitris-rizopoulos. ISBN 9781439872864.